

RESEARCH PAPER

Identification of quantitative trait loci controlling fibre length and lignin content in *Arabidopsis thaliana* stems

Arnaud Capron¹, Xue Feng Chang², Hardy Hall³, Brian Ellis³, Rodger P. Beatson² and Thomas Berleth^{1,*}

¹ University of Toronto-CSB, 25 Willcocks Street, Toronto, ON, Canada, M5S 3B2

² British Columbia Institute of Technology, 3700 Willingdon Avenue, Burnaby, BC, Canada, V5G 3H2

³ University of British Columbia – Michael Smith Laboratories, #301–2185 East Mall, Vancouver, BC, V6T 1Z4, Canada

* To whom correspondence should be addressed. E-mail: thomas.berleth@utoronto.ca

Received 3 April 2012; Revised 11 October 2012; Accepted 15 October 2012

Abstract

Fibre properties and the biochemical composition of cell walls are important traits in many applications. For example, the lengths of fibres define the strength and quality of paper, and lignin content is a critical parameter for the use of biomass in biofuel production. Identifying genes controlling these traits is comparatively difficult in woody species, because of long generation times and limited amenability to high-resolution genetic mapping. To address this problem, this study mapped quantitative trait loci (QTLs) defining fibre length and lignin content in the *Arabidopsis* recombinant inbred line population Col-4×Ler-0. Adapting high-throughput phenotyping techniques for both traits for measurements in *Arabidopsis* inflorescence stems identified significant QTLs for fibre length on chromosomes 2 and 5, as well as one significant QTL affecting lignin content on chromosome 2. For fibre length, total variation within the population was 208% higher than between parental lines and the identified QTLs explained 50.58% of the observed variation. For lignin content, the values were 261 and 26.51%, respectively. Bioinformatics analysis of the associated intervals identified a number of candidate genes for fibre length and lignin content. This study demonstrates that molecular mapping of QTLs pertaining to wood and fibre properties is possible in *Arabidopsis*, which substantially broadens the use of *Arabidopsis* as a model species for the functional characterization of plant genes.

Key words: *Arabidopsis*, composite interval mapping, fibre, lignin, high-throughput, QTL.

Introduction

Wood properties are of great significance to the multitude of nations that produce pulp and paper products. Variation in pulp properties, notably fibre morphology and physical properties such as fibre length, width, strength, and coarseness, have a profound influence on the properties of paper products (Dinwoodie, 1965; Page and Seth, 1980). In woody plants used for pulp production, the traditional method of improving fibre quality has been through tree breeding (Via *et al.*, 2004). Recent work has shown that woody plant fibre cell dimensions and strength are under genetic control (Yu *et al.*, 2001). Lignin,

a major chemical component of wood, also plays an important role in defining the properties of pulp-based products. The lignin content of pulp fibres has to be carefully controlled in order to balance paper strength and optical qualities, and removal of lignin in the pulping process is both costly and environmentally detrimental (Odendahl, 1994; Biermann, 1996). Moreover, lignin content is also a critical parameter in the generation of biofuels from cellulosic materials (Mansfield *et al.*, 2012). Cellulosic biomass, for example from woody species, has become increasingly important as a source of bioenergy

Abbreviations: BAC, bacterial artificial chromosome; BLUP, best linear unbiased predictor; CIM, composite interval mapping; FQA, fibre quality analyser; QTL, quantitative trait locus; RIL, recombinant inbred line; SFP, single-feature polymorphisms.

© 2012 The Authors.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/2.0/uk/>) which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

that is not in competition with food production (Somerville *et al.*, 2010).

Genetic engineering has demonstrated that direct gene manipulation can improve biomass production or wood properties in woody species (Eriksson *et al.*, 2000; Pilate *et al.*, 2002; Ranocha *et al.*, 2002; Bjurhager *et al.*, 2010; Li *et al.*, 2011; Zawaski *et al.*, 2011), but due to strong resistance in the marketplace to the use of genetically modified organisms, it would be preferable to obtain similar results through traditional breeding strategies.

Long generation times, variable growth conditions, and comparatively scarce genomic resources have limited the use of high-resolution genetic mapping in tree breeding, although each of these limitations is addressed in emerging tree genomic systems, such as *Populus trichocarpa* (Tuskan *et al.*, 2006; Jansson and Douglas, 2007). A genetic model plant such as *Arabidopsis thaliana*, on the other hand, is well suited for genetic studies involving large numbers of individuals over multiple generations, because of its small size and short life cycle. Furthermore, a large number of *Arabidopsis* recombinant inbred line (RIL) populations are readily available from stock centres and some of them have been thoroughly characterized genetically (Reiter *et al.*, 1992; Lister and Dean, 1993). *Arabidopsis* RIL populations have been used for fast and accurate identification of numerous quantitative trait loci (QTLs) controlling traits such as circadian rhythm (Swarup *et al.*, 1999), the effects of various nutrients and minerals (Bentsink *et al.*, 2003; Harada *et al.*, 2004; Harada and Leigh, 2006; Reymond *et al.*, 2006; Waters and Grusak, 2008; Zeng *et al.*, 2008; Ghandilyan *et al.*, 2009), resistance to predators (Pfalz *et al.*, 2007), stress responses (McKay *et al.*, 2008), biomass accumulation (Lisec *et al.*, 2008), heterosis (Lisec *et al.*, 2009), and salicylic acid pathway responses (Alcazar *et al.*, 2009).

The initial identification of QTLs in *Arabidopsis* has enabled subsequent detailed genetic studies to unravel the molecular mechanisms controlling the traits associated with the QTLs, and these insights have often contributed to a better understanding of angiosperm plant biology in general (El-Assal *et al.*, 2001; Kroymann *et al.*, 2003; Mouchel *et al.*, 2004; Masle *et al.*, 2005; Werner *et al.*, 2005; Bentsink *et al.*, 2006; El-Din Zhang *et al.*, 2006).

Here, we have explored the feasibility of detecting QTLs in an *Arabidopsis* RIL population for the traits of fibre length and lignin content in inflorescence stems. We have described the use of high-throughput phenotyping protocols for both traits and the molecular mapping of significant QTLs controlling fibre properties and lignin content in any plant species.

Materials and methods

Plant material

A set of 98 RILs derived from a cross between Columbia (Col-4) and Landsberg *erecta* (Ler-0) (Lister and Dean, 1993) was obtained from the ABRC (Stock number CS1899). Seeds were surface sterilized, plated on 0.5× MS agar medium (2.1 g l⁻¹ MS salts, 10 g l⁻¹ sucrose, 8 g l⁻¹ agar, pH 5.8) and vernalized for 48 h at 4 °C. Seedlings were transferred to Professional Pro-Mix 'BX'/Mycorise Pro (Premier Horticulture, Rivière du Loup, Canada) supplemented with 0.3% (v/v) Nutricote 14-14-14 Type 100 fertilizer (Chisso-Asahi Fertilizer Co., Tokyo, Japan) after 10 d of continuous light. Plants were grown under a 16 h/8 h day/night cycle at 21 °C day and 18 °C night temperature. Light was maintained at

200 µmol s⁻² m⁻². Post-flowering inflorescence stems were collected for fibre quality analyser (FQA) processing or lignin content measurement. The set of lines was grown five times, in five separate experiments, and the resulting stems were analysed separately. Three sets of plants were used for fibre length measurements and two sets of plants were used for lignin content measurements. A single plant per line was used in the first FQA experiment and two plants per line were bulked in all subsequent experiments.

Isolation of fibres from plant stems

Samples (10 mg) of air-dried stem material taken from the bottom 5 cm of one stem or each of two stems, node and internode, were placed in a 20 ml test tube for the pulping reaction. The stem sample was compacted using a glass rod before adding 2 ml each of distilled water and acetic acid (glacial) and heating in a boiling water bath for 2 min. After adding 2 ml of 30% hydrogen peroxide, the samples were returned to the boiling water bath for 90 min. The resulting solution was then carefully decanted to retain the cooked stem tissue within the tube, and the tissue was rinsed gently with distilled water three times to remove residual reagents. The delignified stem tissue was transferred to a screw-cap conical plastic centrifuge tube with 35 ml of distilled water and then agitated vigorously to disintegrate fibre bundles and form a homogenous fibre suspension. The suspension was subsequently filtered through a Britt Dynamic Drainage Jar using 3 l of water, stirred using an overhead stirrer at 200 r.p.m. (TAPPI, 1992) to collect fibres retained on the 200-mesh screen (105 µm opening). The retained fibres were rinsed off the filter mesh, collected in a 50 ml centrifugal tube and diluted to a total volume of 50 ml with distilled water. The fibre suspensions were inspected visually and fibre bundles, if present, were removed manually using a fine needle. Triplicate measurements were performed on 10 mg samples taken from individual stems of the same line.

Fibre quality analyser measurements

Fibre length was determined using a FQA (OpTest; Hawkesbury, ON, Canada) with a cytometric flow cell and image analysis system capable of rapidly and accurately measuring fibre curl, kink, and length distributions (Olson *et al.*, 1995; Roberston *et al.*, 1999). Five millilitres of the fibre suspension was dispensed into a 600 ml plastic beaker for FQA analysis. This sample was then diluted automatically to exactly 600 ml by the FQA. The fibre input was adjusted to a targeted events s⁻¹ measurement range of 25–40 fibres s⁻¹. The length weighted fibre length is reported.

Removal of extractives

The inflorescence stems of *Arabidopsis* accessions were first ground using a microball mill to pass through a 80-mesh screen. Prior to extraction, the ground samples were dried in a vacuum oven at 40 °C for 48 h and conditioned in a vacuumed desiccator over phosphorus pentoxide overnight. The ground stems were then extracted using a rapid extraction by washing method. The procedure was based on Morrison's quick extraction method (Morrison, 1972) with a reduction in sample particle size and weight, and passage through a fine filtration membrane. Approximately 0.1 g of dried 80-mesh sample was put in a test tube, soaked with 15 ml of distilled water, heated in a water bath at 65 °C for 30 min with occasional shaking, and then filtered hot through a dry and pre-weighed 0.45 µm nylon membrane using a Millipore filter. The residue was first washed with 20×2 ml of deionized water and then sequentially with 20×1 ml each of ethanol, acetone, and diethyl ether. The residue was then transferred to a pre-weighed aluminium pan in preparation for lignin content measurement.

Lignin content measurement

The procedure used was as described by Chang *et al.* (2008). Dried 80-mesh samples (5±1 mg) of the ground and extracted *Arabidopsis* stems were weighed to the nearest 0.01 mg, and then digested with 1.0 ml of 25% acetyl bromide in acetic acid in a 70 °C water bath for

30 min with shaking at 10 min intervals. After cooling at room temperature, the samples were stored in an ice bath for 5 to 120 min, and 5.0 ml of acetic acid was added to each sample. After mixing, 30 µl aliquots of the mixed content, in triplicate, were transferred to the wells of a 96-well quartz microplate. Transfer of the samples to the microplate was completed within 5 min before sequentially adding 40 µl of 1.5 M NaOH, 30 µl of 0.5 M hydroxylamine hydrochloride and 150 µl of acetic acid to each well using a ten-channel multiple pipette. The absorbance of the solutions in the wells at 280 nm was measured using a Perkin-Elmer Wallac 1420 microplate reader. A blank was included to correct for background absorbance by the reagents.

Data analysis

Data analysis was carried out using the R statistical language (R Development Core Team, 2009). Broad sense heritability was estimated using the analysis of variance component in the lme4 package. Variance components for the lines and for the experiments were calculated with the model:

$$\text{lmer}(\text{TRAIT} \sim (1|\text{LINE}) + (1|\text{EXPERIMENT}))$$

and used to calculate the heritability with the formula:

$$H^2 = \text{Var}(\text{lines}) / [\text{Var}(\text{lines}) + \text{Var}(\text{Residual})]$$

Best linear unbiased predictor (BLUP) values for each line were also calculated with the lme4 package. Variance components for the lines and the experiments were calculated using the same model as for the heritability, and the random effects for each line were extracted with the ranef function.

A composite interval method was performed with the R/qtl (Broman *et al.*, 2003) analysis package. The imputation method was selected, with 256 draws, an error of 0.001, and 2.5 cM steps. The composite interval mapping was done with the imputation method, with three covariables and a 10 cM window:

$$\text{cim}(\text{cross}, \text{method} = \text{'imp'}, \text{window} = 10, \text{n.marcover} = 3)$$

Calculations were performed on each experimental data set independently and also on the calculated BLUP values for fibre length and lignin content using the same method. The significant thresholds were determined by performing 1000 permutations. The multiple QTL model was devised using the fitqtl function. Results obtained from the composite interval mapping analysis were used to construct the QTLs and their positions were used in a simple additive model. The resulting model was used to assess the effect of each QTL on the trait (percentage of explanation of the observed total variance).

Analysis of variance (ANOVA) was performed to test the association of the genotype of selected markers with the observed trait, using the gee and multcomp packages.

Candidate gene selection

The list of all protein-encoding genes included in the intervals delimited by FQ2 and FQ5 (see Results) was screened for genes with enriched expression in the second-internode data set of the Bio-Array Resources for Plant Biology's expression browser (<http://bar.utoronto.ca/welcome.htm>). The selected cut-off was 1.5 on the log₂ scale. Additional genes were also retained for their known connection to fibre and cell-wall development or auxin/gibberellin connection. The selected list of genes was then screened for mismatches at the amino acid level between the resequenced Col-0 genome (Cao *et al.*, 2011) and the Ler-0 genome (Gan *et al.*, 2011), using the Blastx algorithm (Altschul *et al.*, 1997). Genes showing at least one change at the amino acid level were considered as candidates. For the LQ2 interval (see Results), the same method was applied, but additional genes were limited to lignin metabolism/synthesis.

Results

Fibre length measurements

The Col-4×Ler-0 RIL population created by Lister and Dean (1993) was chosen to produce the stem tissues used for the analysis. To collect material for fibre length measurements, the RIL population was grown in three independent biological samples. A single, central inflorescence stem from each line was used in the first experiment, referred to as Fibre1, whilst two stems from two individuals per line were pooled together in the two subsequent experiments, referred as Fibre2 and Fibre3. Great care was taken to minimize contamination from other cell types, typically leading to a very clean fibre preparation (data not shown). The FQA analysis yielded considerably higher fibre length values in the Col-4 parental line than in the Ler-0 parental line, as shown by a *t*-test ($P < 0.001$), the mean value across the three experiments being 0.977 and 0.757 mm, respectively.

The fibre length values determined across the entire set of RILs showed a wider range than the values observed for the parental lines, which indicated the possibility of transgressive segregation (Fig. 1). The shortest measured fibres within the RILs were 0.621 mm long, 83.9% of the minimum Ler value, whilst the longest were 1.287 mm, 122.5% of the longest Col value (Table 1). Broad sense heritability for the trait was estimated to be 62.3%. The histograms from the individual experiments for Fibre1, Fibre2, and Fibre3 (Fig. 1A, B, and 1C, respectively) showed a similar structure. In each, a main peak was associated with shorter fibres, as indicated by the median values in Table 1, with a trailing end towards the longer values.

Lignin content measurements

For lignin content measurements, the RIL population was grown twice to harvest material for biological replicas, labelled Lignin1 and Lignin2. Measured differences in lignin content between the parent lines were inconclusive; a 2% difference was seen between the parents in Lignin1 but almost none in Lignin2 (Table 2). This was confirmed by a *t*-test, which did not show a statistically significant difference between the parents ($P > 0.60$). The mean value of both experiments was a lignin content of 19.65% (w/w) (Table 2).

In contrast, the mean lignin content values in individual lines of the RIL population ranged from 13.13 to 20.98%. Broad-sense heritability was estimated at 32.3%, possibly indicating a relatively limited genetic influence. The distribution of lignin content values in the Lignin1 and Lignin2 datasets was very similar, as shown in Figs 2A and B.

QTL analysis for fibre length

A genetic map based on the segregation of 676 single-feature polymorphisms (SNPs) in the Col-4×Ler-0 RIL population (Singer *et al.*, 2006) was used to perform a QTL analysis using the fibre length values of 98 lines of the RIL collection, first by interval mapping (data not shown) and then by composite interval mapping (CIM) on each data set independently, using R/qtl (Jansen, 1993; Zeng, 1993, 1994; Jansen and Stam, 1994).

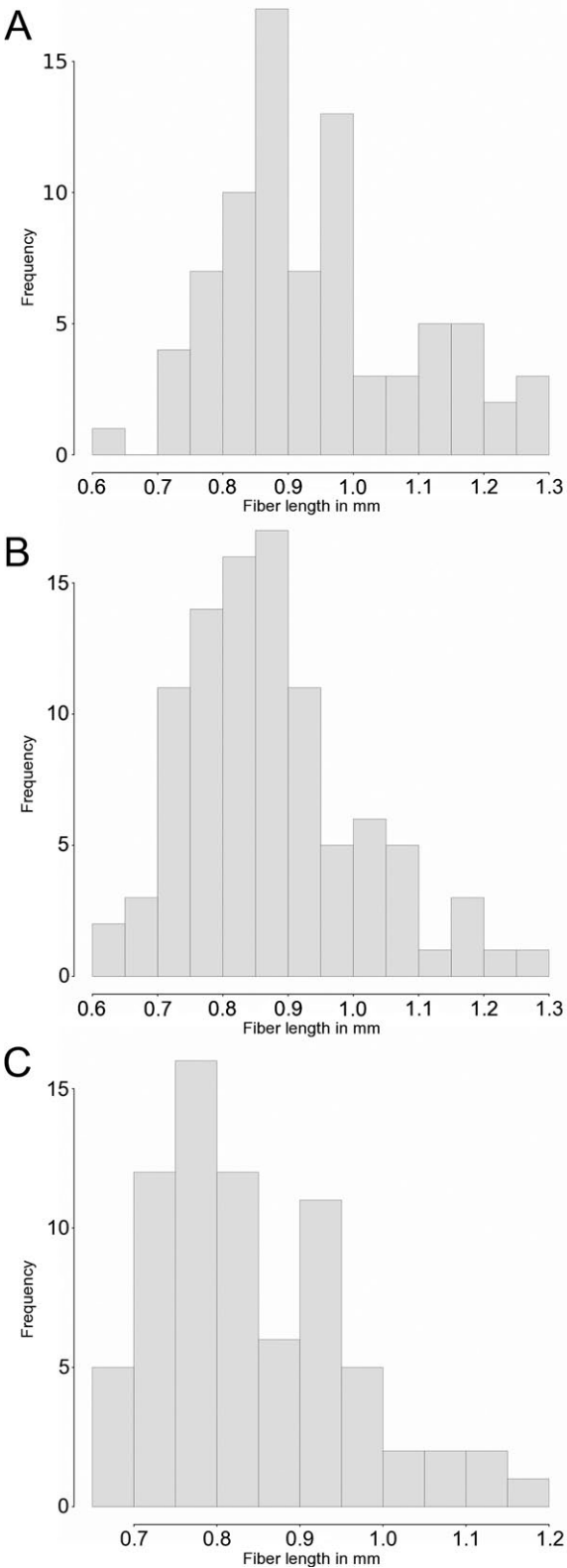


Fig. 1. Distribution of fibre length values (mm) in the three experiments: Fibre1 (A), Fibre2 (B), and Fibre3 (C).

Two significant fibre length QTLs ($P < 0.05$) were mapped to similar genome positions in all three data sets (Fibre1, Fibre2, and Fibre3): based on the chromosome to which they each mapped, they are referred to hereafter as FQ2 and FQ5, respectively.

Table 1. Fibre length measurements (mm) in the RILs and Col and Ler parents.

Experiment	Parents	RIL				
		Fibre length	Mean	Mean	Median	Range SD
Fibre1	Col	1.051	0.926	0.943	0.909	0.643–1.287
	Ler	0.800				
Fibre2	Col	0.935	0.837	0.873	0.852	0.621–1.267
	Ler	0.740				
Fibre3	Col	0.945	0.839	0.840	0.817	0.650–1.181
	Ler	0.732				

Table 2. Lignin content measurements in the RILs and Col and Ler parents All values are given as the percentage of lignin based on unextracted weight (w/w).

Experiment	Parents	RIL				
		Lignin content	Mean	Mean	Median	Range SD
Lignin1	Col	17.74	18.73	17.09	17.15	13.13–19.39
	Ler	19.73				
Lignin2	Col	20.40	20.57	18.07	18.09	13.80–20.98
	Ler	20.75				

FQ2 was located on the bottom arm of chromosome 2, with the maximum LOD score positioned between 49.5 and 55 cM (Fig. 3) in the map defined by Singer *et al.* (2006). The maximum LOD score was observed in the Fibre2 data set at 9.09, and LOD scores in Fibre1 and Fibre3 were also highly significant, with LOD scores of 8.63 and 5.67, respectively (Fig. 3). The extent of the QTL varied slightly between the data sets, but the overall confidence interval could be calculated based on a 2 LOD score drop from the summit of the peak. This placed FQ2 between 43.3 and 57 cM, an interval covering 21 SFP markers (Fig. 3, Table 3).

The second significant QTL, FQ5, was located in the distal part of the top arm of chromosome 5. The maximum LOD score was found between 3 and 8 cM, with a maximum value of 9.90 in the Fibre1 experiment (Fig. 3). The estimated extent of FQ5 with a 2 LOD score drop was between 1.5 and 11.6 cM, which encompassed 17 SFP markers (Table 3).

In order to assess the accuracy of the mapping, BLUP values were calculated from the three experiments and used to perform the CIM calculation, producing results very similar

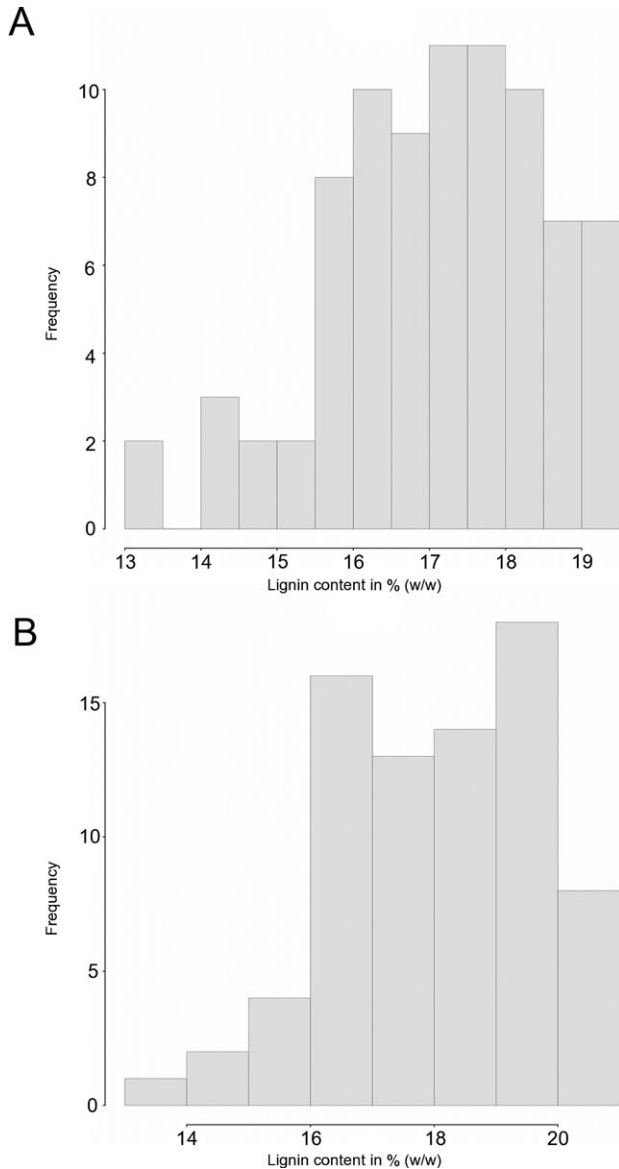


Fig. 2. Distribution of the values for lignin content (% w/w) in the two experiments: Lignin1 (A) and Lignin2 (B).

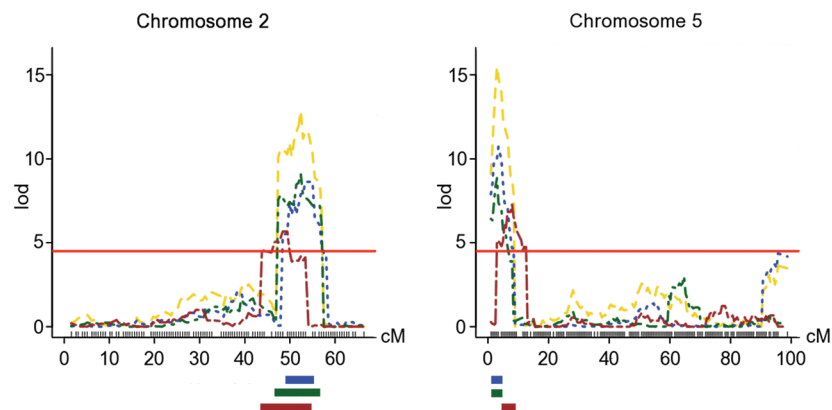


Fig. 3. Distribution over chromosomes 2 and 5 of the LOD scores for fibre length. LOD was calculated independently by CIM for the three fibre experiments using a 10 cM window and three covariates. The LOD score for Fibre1 is in blue, Fibre2 in green, Fibre3 in dark red, and BLUP in yellow. The red line shows the most conservative statistically significant threshold ($P < 0.05$), calculated for Fibre1 as 4.35. Under each graph is presented the QTL interval corresponding to a 2 LOD score drop.

to those obtained with each individual experiment, as shown in Fig. 3.

Fig. 4 showed that the measured fibre length was strongly correlated with the genotype of two markers (C2_081 and C5_008) in all three experiments, a fact further confirmed by ANOVA (Supplementary Table 1 at *JXB* online). These markers were selected as proxies for FQ2 and FQ5, as they were both found as part of the interval identified in all three Fibre experiments. Plants having the Ler allele of FQ2 and the Col allele of FQ5 displayed the largest fibres (mean 0.947–1.086 mm), whilst the opposite allele combination yielded the shortest fibres (mean 0.764–0.821 mm). Fibre lengths in individuals possessing purely Col or Ler alleles at these sites fell in between these extremes.

Further analysis with the R/qtl package using a multiple QTL model (Haley and Knott, 1992; Sen and Churchill, 2001) was performed to estimate the contribution of each QTL to the observed variance, based on a model devised for each data set. In the three data sets, the contribution of FQ5 to the total variance (23.96–29.96%) was consistently higher than that of FQ2 (15.91–22.18%). Importantly, the estimated combined contribution of both QTLs amounted to approximately one-half of the total variance (45.79–50.58% depending on the data set).

QTL analysis for lignin content

The genetic map by Singer *et al.* (2006) was also used to locate QTLs controlling lignin content. Interval mapping and CIM revealed a strong LOD score peak on chromosome 2 in both independent experiments, indicative of a significant QTL, henceforth referred as LQ2 (Fig. 5). The margins, defined by a 2 LOD score drop, located the QTL between 35.3 and 47.9 cM, an interval containing 23 SFP markers. As for the fibre experiment, BLUP values were calculated from both lignin experiments and used to perform the same CIM calculation. The results obtained with the BLUPs were very similar to those seen with individual experiments (Fig. 5). A multiple QTL model was again used to determine the contribution of the significant QTL to the observed variance (Table 4), and this revealed that between 24.80 and 26.51% of the measured variance in lignin content was

Table 3. Summary of the QTLs identified by composite interval mapping. The positions and marker intervals are based on the map of [Singer *et al.* \(2006\)](#).

QTL	Experiment	Chromosome	Marker interval	Position (cM)	LOD
Fibre length					
FQ2	Fibre1	2	C2_079-C2_089	53.4	8.63
	Fibre2	2	C2_075-C2_091	52.4	9.09
	Fibre3	2	C2_071-C2_088	49.1	5.67
FQ5	Fibre1	5	C5_002-C5_008	3.5	10.07
	Fibre2	5	C5_002-C5_008	3.0	8.85
	Fibre3	5	C5_008-C5_018	8.0	7.27
Lignin content					
LQ2	Lignin1	2	C2_067-C2_077	45.9	5.37
	Lignin2	2	C2_056-C2_072	40.8	5.84

attributable to LQ2 in the two independent experiments. Sorting the individual lines based on a single marker (C2_067) showed that individuals carrying the Ler allele of this marker had, on average, less lignin (16.50–17.47%) than individuals carrying the Col allele (17.92–19.00%) ([Fig. 6](#)), a fact also confirmed by ANOVA ([Supplementary Table 1B](#)). The marker C2_067 was selected as a proxy for the QTL LQ2, as it belonged to the interval defined in both Lignin experiments.

Annotated genes within the QTL intervals

As all the SFP markers used in the construction of the [Singer *et al.* \(2006\)](#) genetic map correspond to positions in the fully sequenced genome of *Arabidopsis*, all annotated genes within the QTL intervals identified by CIM could readily be identified. The QTL FQ2 was flanked by the markers C2_071 and C2_091, which corresponded to the genes *At2g28060* and *At2g37050* according to the AGI version 10 of the *Arabidopsis* genome. This interval is 3.6 Mb long and contains 1005 loci encoding proteins. The second significant fibre length QTL FQ5, flanked by C5_002 and C5_018, was located in a 2.7 Mb interval between *At5g01150* and *At5g08580*, which contains 789 protein-encoding genes. Finally, LQ2, the QTL controlling lignin content, was located between the markers C2_056 and C2_077, corresponding to *At2g23900* and *At2g30240*, a 1.12 Mb interval that contains 672 protein-encoding loci.

Although our genetic approach could link entirely uncharacterized genes to fibre development, it is be worthwhile mining the identified intervals for plausible candidate genes, as the use of increasing genomic resources could further narrow down the selection and accelerate the cloning of the respective QTLs. On the basis of currently available resources, we selected candidate genes in the following way. First, genes with an association with fibre and cell-wall development were compiled. This list was amended for genes involved in auxin or gibberellin metabolism and signalling, two hormones with roles in fibre development ([Zhong and Ye, 2001](#); [Dayan *et al.*, 2010](#); [Ragni *et al.*, 2011](#)). At the same time, the Bio-Array Resources for Plant Biology's expression browser ([Toufighi *et al.*, 2005](#)) was used to identify genes whose expression was enriched in the stem

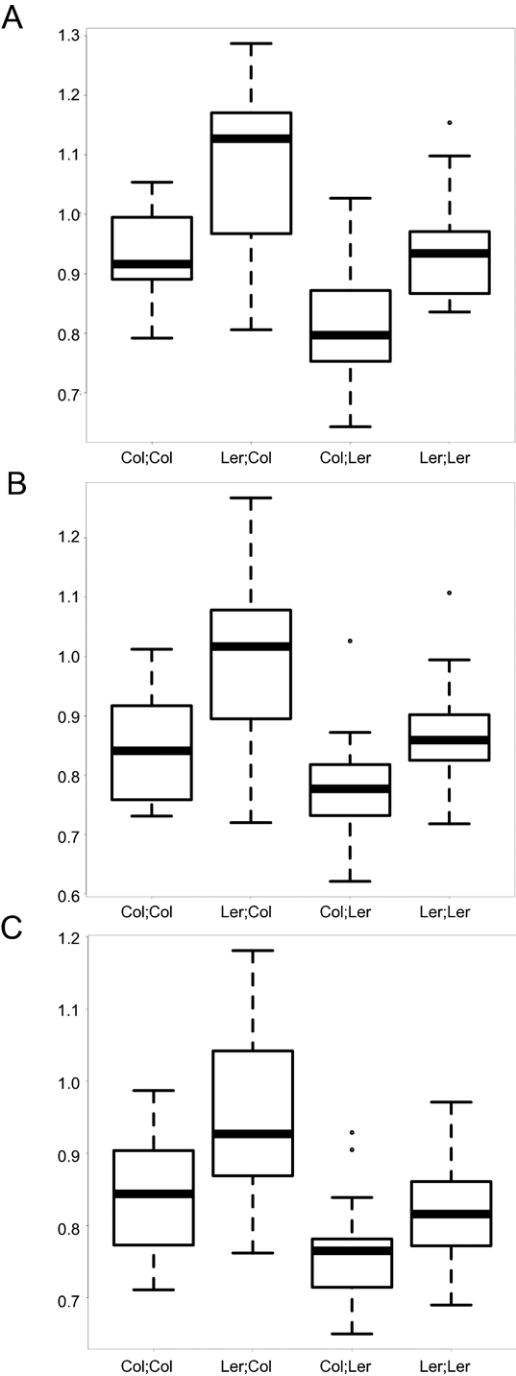


Fig. 4. Box plots of fibre length sorted by two markers each linked to QTLs for the Fibre1 (A), Fibre2 (B), and Fibre3 (C) experiments. The box plots were drawn after sorting the population according to the genotype of the markers C2_081 and C5_008, showing the difference in average fibre length (mm) of each combination. The graphs clearly show that individuals with the combination Ler C2_082 and Col C5_008 had the longest fibres, whilst individuals with the combination Col C2_082 and Ler C5_008 had the shortest fibres. The x-axis shows the genotype of the markers C2_082 and C5_008 and the y-axis shows fibre length.

second-internode dataset. Finally, all the selected genes were screened for predicted amino acid polymorphisms between the Columbia and Landsberg *erecta* alleles. This selection resulted

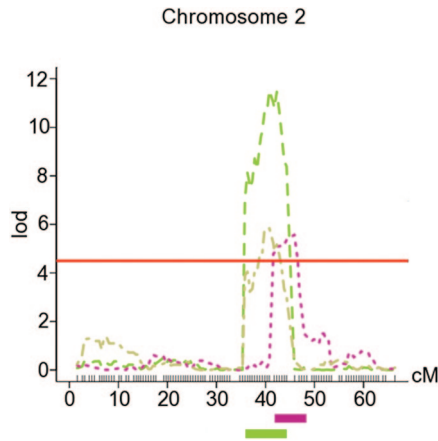


Fig. 5. Distribution over chromosome 2 of the LOD score for lignin content. LOD was calculated independently by CIM for the two lignin experiments using a 10 cM window and three covariates. The LOD score for Lignin1 is in green, Lignin2 in pink, and the BLUP in khaki. The red line shows the most conservative statistically significant threshold ($P < 0.05$), calculated for Lignin1 as 4.39. Under the LOD score curve are indicated the intervals obtained from Lignin1 and Lignin2 in their respective colours.

Table 4. Summary of the QTL in the multiple QTL mapping model. QTLs are identified by their location. The % variance is the percentage of the total variance explained by the QTL. The LOD score given is for the entire model.

QTL	Experiment	% Variance	<i>P</i> value	LOD score
Fibre length				
FQ2	Fibre1	22.18	3.33 ^{e-08}	13.76
FQ5	Fibre1	28.40	1.03 ^{e-09}	
FQ2	Fibre2	21.83	3.39 ^{e-09}	15.46
FQ5	Fibre2	23.96	8.03 ^{e-10}	
FQ2	Fibre3	15.91	6.87 ^{e-06}	11.83
FQ5	Fibre3	29.96	4.80 ^{e-09}	
Lignin content				
LQ2	Lignin1	26.51	7.45 ^{e-07}	5.49
LQ2	Lignin2	24.80	4.71 ^{e-06}	4.70

in 51 genes for FQ2 and 27 for FQ5. A complete list is provided in [Supplementary Table 2](#) (at JXB online).

Some of the selected genes were highly conspicuous. For example, a gene involved in the metabolism or deposition of cell-wall polymers other than cellulose, *FRAGILE FIBER 8/IRREGULAR XYLEM 7 (FRA8/IRX7)* encoding a glycosyltransferase involved in the synthesis of xylan, could affect fibre length ([Brown et al., 2007](#)). Remarkably, the *fra8* mutant also displays thinner secondary cell walls in fibres ([Zhong et al., 2005](#)). Among the genes found in the FQ2 interval, there were several members of the *CELLULOSE SYNTHASE-LIKE (CSL)* family: *CSLA7*, *CSLB1*, *-B2*, *-B3*, and *-B4*, and *CSLD1*. *CSLD1* is a gene whose activity is important for pollen-tube cell-wall development and growth ([Bernal et al., 2008](#)), while *CSLA7* encodes a protein reportedly involved in the production β -mannans ([Liepman et al., 2005](#)). The function of β -mannans within the cell wall remains unknown,

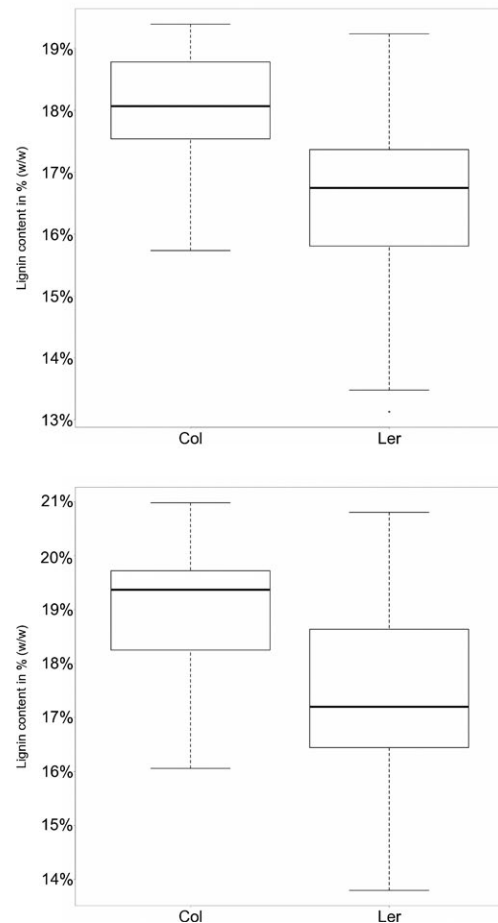


Fig. 6. Box plots of the lignin content values sorted according to a QTL-linked marker for the Lignin1 (A) and Lignin2 (B) experiments. The box plots were drawn after sorting the population according to the genotype of the marker C2_067, showing the difference in average lignin content. The graphs clearly show that individuals with the Ler marker had the lowest lignin content.

but *CSLA7* activity is crucial for proper embryo development. A *csla7* null allele causes arrest during embryogenesis, disrupting normal cell division and patterning of the embryo, precluding further analysis of the protein function at later stages ([Goubet et al., 2003](#)). However, the precise functions of the *CSLB* genes have not been investigated. Also, genes involved directly in cellulose synthesis deposition or related processes, whose knockout mutant phenotypes may affect many tissues, may quantitatively influence fibre properties in their allelic variation, as found in FQ2 and FQ5. The *TRICHOME BIREFRINGENCE (TBR)* gene is located inside FQ5. *TBR* is involved in synthesis of the secondary cell wall, as is its homolog *TBR-LIKE 3 (TBL3)*, which is also present in FQ5. Mutant *tbr* and *tbl3* plants display reductions in crystalline cellulose levels but not in total cellulose content ([Bischoff et al., 2010](#)). Furthermore, more members of the *TBL* gene family were found in FQ5 (*TBL35*) and FQ2 (*TBL43*, 45). In addition, lignins play major roles in determining the mechanical properties of the cell wall, and mutants affected in lignin biosynthesis or deposition may display fibre abnormalities. Among the candidate genes identified were four laccase-encoding genes: one in FQ2, *LAC2*, and three on FQ5, *LAC10*,

-11, and -12. Laccases are ubiquitous copper oxidases, present in fungi and plants (Messerschmidt and Huber, 1990) and have been used for their lignin-degrading activity in fibre pulping (Mayer and Staples, 2002). In *Arabidopsis*, the laccase gene family comprises 17 members, of which only two have been subjected to detailed study: *LAC4* and -17. Mutations in these genes result in reduced general lignification, a phenotype particularly pronounced in the interfascicular fibres in the *lac17* mutant (Berthet *et al.*, 2011). Another gene involved in lignin metabolism in FQ2 is the cinnamate 4-hydroxylase (C4H) *At2g30490*. A mutation in this gene, *reduced epidermal fluorescence 3 (ref3)*, results in collapsed stem vasculature and reduced lignin content (Schillmiller *et al.*, 2009).

A link between fibre defects and auxin transport defects is plausible in the light of cell-dimension changes in auxin-transport-inhibited plants (Mattsson *et al.*, 1999) and has recently been supported by the *walls are thin 1 (wat1)* mutant in *Arabidopsis* (Ranocha *et al.*, 2010), as well as by a mutation in the *LIKE AUXIN RESISTANT1 (LAX1)* gene, which is located in FQ5 and encodes an auxin influx carrier of the AUX1 family (for review, see Kramer, 2004). Moreover, the *ATP BINDING CASSETTE B1* gene encoding an ABC transporter, also part of an auxin transport system, is located in the FQ2 interval (Wu *et al.*, 2010). Gibberellins are also known to influence cell dimensions and Dayan *et al.* (2010) demonstrated a role of gibberellins in the development of fibres through modulation of GA 2-oxidase and GA 20-oxidase activity. These findings prioritize GA 2-oxidase in FQ2 (GA2OX3) and GA 20-oxidase in FQ5 (GA20OX3) as candidate genes for further studies.

Other genes deserve close attention because of their related mutant phenotypes. The *SCAR/WAVE 1* encodes an activator of the ARP2/3 complex responsible for the nucleation of the actin filaments and, together with the other *SCAR* genes, is necessary for normal stem growth (Zhang *et al.*, 2008). *BELLRINGER/PENNYWISE (BLR/PNY)* encodes a BELL1-like transcription factor directly involved in the patterning of the inflorescence stem. A *brl/pny* mutant shows defects in shoot development and displays an altered stem anatomy (Byrne *et al.*, 2003; Smith and Hake, 2003). *BLR/PNY* can control internode cell elongation through modulation of the activity of the pectin methylesterase PME5 (Peaucelle *et al.*, 2011). Interestingly, *At2g36700* and *At2g367210* encode a pair of pectin methyl-esterases and *At5g04970* encodes a pectin methylesterase inhibitor, although no particular phenotypic functions are known for them. *FOLYL POLYGLUTAMATE SYNTHETASE 1 (FPGS1)* is a plastidial enzyme belonging to a small gene family. Whilst the *fpgs1* mutant has no discernable phenotype, the double mutant of *fpgs1* and *fpgs3* is dwarfed, indicating a role in stem growth (Mehrrshahi *et al.*, 2010).

Finally, another group of genes located in the FQ5 interval might also affect fibre properties through their roles in cell proliferation or cell division. One of these is the *CELL-DIVISION CYCLE 48C (CDC48C)* gene, one of three *CDC48* genes in *Arabidopsis*. The roles of *CDC48* genes in cell division and cytokinesis have been characterized in detail (Rancour *et al.*, 2002; Park *et al.*, 2008). Furthermore, the SIAMESE (SIM) protein has a cyclin-binding motif and interacts with several cell-cycle regulators (Kasili *et al.*, 2010; Van Leene *et al.*, 2010). *SIM*

controls the endoreduplication in trichomes and other organs (Walker *et al.*, 2000).

Eighteen genes from the LQ2 interval were selected as candidate genes because of suspected or demonstrated impact on lignin biosynthesis or by their enriched expression in the stem second-internode in association with predicted amino acid polymorphisms between Columbia and Landsberg erecta. Among these, the *CCR6* gene, *At2g23910*, encodes a putative cinnamoyl CoA reductase, a member of a family of oxidoreductases that includes enzymes responsible for the final step in monolignol biosynthesis (for review, see Boerjan *et al.*, 2003). Studies in tobacco and eucalyptus have linked CCRs and lignin biosynthesis (Lacombe *et al.*, 1997; Ralph *et al.*, 1998). Although the precise function of *CCR6* is unknown, it is most strongly expressed in immature seeds, and microarray surveys have shown it to be co-expressed with flavonol biosynthesis genes (Yonekura-Sakakibara *et al.*, 2008). Among the 11 *CCR* genes annotated in *Arabidopsis*, only *CCR1 (At1g15950)* has been characterized extensively. A *crr1* mutant seems to suffer from reduced lignin content in the stem (Jones *et al.*, 2001), although more recent work suggests that this phenotype may be the result of delays in lignin deposition rather than an inability to execute the reaction (Patten *et al.*, 2005; Laskar *et al.*, 2006).

Discussion

Whilst QTLs affecting fibre properties have been mapped in several plant species, including maize (Cardinal *et al.*, 2003; Krakowsky *et al.*, 2006, 2005), sorghum (Shiringani and Friedt, 2011), and, most notably, cotton (Mei *et al.*, 2004; Chee *et al.*, 2005a,b; Draye *et al.*, 2005; Ulloa *et al.*, 2005; Qin *et al.*, 2008; Chen *et al.*, 2009; Paterson *et al.*, 2011; Zhang *et al.*, 2011), application of similar approaches to trees has been hindered by the length of time required for the tree to mature, combined with the difficulty in predicting final wood properties in mature trees based on the properties of juvenile wood.

Although QTL mapping in crop plants and trees directly benefits breeding efforts, molecular access to relevant genes can be achieved much faster in genetic model plants, such as *Arabidopsis*, where biological properties and genomic resources enable unmatched genetic resolution and molecular access to underlying genes, although it can suffer from its strength, as its simplicity might mask complex epistatic interactions found in larger crop genomes. Furthermore, *Arabidopsis* has also demonstrated its suitability for providing molecular access to conserved plant biological mechanisms (Liepman *et al.*, 2010; Wienkoop *et al.*, 2010; Zhang *et al.*, 2011), and, as an arboreal growth habit is not a monophylogenetic trait but has arisen independently many times in both angiosperm and gymnosperm lineages (Bowe *et al.*, 2000; Chaw *et al.*, 2000; Soltis *et al.*, 2002), our understanding of the molecular mechanisms operating in the cells of trees is expected to benefit from genetic discoveries in *Arabidopsis* as much as in other plants.

In this study, we explored the possibility of identifying QTLs relevant for important wood traits in an established RIL population of *Arabidopsis* as a first step towards a molecular understanding of the underlying cellular mechanisms. Our reproducible

mapping of two QTLs controlling fibre length in *Arabidopsis* accessions have demonstrated the feasibility of the approach, as well as of the high-throughput phenotyping technology that was employed in the study. Rapidly developing mass-sequencing and positional cloning technologies are well suited to move on towards the identification of the molecular alleles underlying these two QTLs in the future. At the same time, *Arabidopsis* QTL analysis can be expanded towards accessions with more strongly divergent fibre traits.

Our results for genetic control of lignin content seemed to indicate a limited genetic influence, but it was still sufficient to successfully perform QTL mapping. The results of CIM analysis were similar in the Lignin1 and Lignin2 experiments (Figs 4 and 6), which together identified the first QTL affecting lignin content in the *Arabidopsis* stem. As with fibre properties, studies on genetic control of lignin content have been conducted in other species, such as pine (Sewell *et al.*, 2002), maize (Cardinal *et al.*, 2003; Krakowsky *et al.*, 2006, 2005), barley (Grando *et al.*, 2005) and rice (Xie *et al.*, 2011), research that might be complemented with molecular access to the relevant genes in *Arabidopsis*. As part of our mapping efforts, we have incorporated a novel, micro-scale, acetyl bromide-based method for determination of lignin content (Chang *et al.*, 2008) developed for assaying large numbers of *Arabidopsis* plants. The results reported here encourage its application for screening a wider range of *Arabidopsis* accessions to identify extreme divergences of lignin content in this species. Genetic analysis of such genotypes may in turn lead to the identification of more numerous and stronger QTLs for this trait.

Previous studies have identified a number of factors influencing fibre properties, including hormonal regulation (Zhong and Ye, 2001; Dayan *et al.*, 2010; Ragni *et al.*, 2011), activation of a number of key transcription factors (Ko *et al.*, 2007; Yamaguchi *et al.*, 2011) and cytoskeletal functions (Burk and Ye, 2002). Although mutations in any of these pathways could affect fibre length, and thus generate the spurious impression that in the background of all these known influences no new determinants of fibre length might be found, this is not an uncommon situation for new genetic searches. Such searches, however, regularly find new influences and their underlying molecular bases, which can then be linked to the network of previously known mechanisms.

Whether the QTLs affect genes with a known relationship to fibre properties or entirely new genes, we can exclude that there is any obvious hormonal, cytoskeletal, or other known influence overtly segregating in the Col-4×Ler-0 RIL population that would just impinge on fibre properties. Furthermore, we explored the fidelity of FQA measurements by comparing them with microscopically measured fibre lengths in a large number of *Arabidopsis* accessions (R.P. Chandra, H.X. Chang, G. Soong, A. Capron, T. Berleth, and R.P. Beatson, unpublished data). We also did not observe any correlation with plant height or stem elongation, either in this study or in a systematic survey of 150 *Arabidopsis* accessions (R.P. Chandra, H.X. Chang, G. Soong, A. Capron, T. Berleth, and R.P. Beatson, unpublished data). This is supported by the fact that there is no straightforward correlation between plant height and cell size. For example, mutants defective in fibre development are not necessarily reduced in height (Zhong *et al.*, 1997). Furthermore, although the *erecta* mutation reduces plant height, cortex cells in the mutant have

been described to be longer (Torii *et al.*, 1996). In conclusion, the absence of any recognizable correlation between fibre length and any other trait in the segregating population underscores the need to find relevant genes underlying the QTLs FQ2 and FQ5 to gain a more comprehensive understanding of fibre length regulation.

It is noteworthy that, for both the fibre length and lignin content traits, the variation within the population vastly exceeded the variation between the parental genotypes, a phenomenon that has been observed in other similar studies involving RIL populations (Reymond *et al.*, 2006; Coluccio *et al.*, 2011; Sanyal and Randal Linder, 2011). Multiple QTLs may mask each other's effects within the genome of a parental line but then become fully recognizable in a segregating population. Thus, it is worthwhile searching for QTLs in established RILs first, which provides the advantages of a high density of DNA markers and thorough characterization of the parental genotypes. In the case of Col and Ler, the two most widely used *Arabidopsis* accessions, their fully sequenced genomes and the availability of bacterial artificial chromosome (BAC) libraries are invaluable aids for identifying the molecular basis of each QTL. The ability of one of these BACs (or an overlap of several BACs) to confer altered fibre length properties to another accession would be the entry point for the molecular genetic dissection of fibre length control.

Association genomics studies are presently underway in woody species such as *Populus* and *Eucalyptus*, whose genomes have been fully sequenced. These analyses seek to correlate single-nucleotide polymorphism allele patterns with variance in a wide range of wood properties, including fibre length and lignin content (reviewed by Nieminen *et al.*, 2012; Mizrachi *et al.*, 2012). As gene sequence and overall synteny are largely conserved between *Arabidopsis* and *Populus* (Tuskan *et al.*, 2006), for example, it will be interesting to compare the results of such genome-wide analyses with the gene content of the *Arabidopsis* QTL intervals detected here.

Supplementary data

Supplementary data are available at *JXB* online.

Supplementary Table S1. ANOVA testing for correlation between traits and selected markers.

Supplemental Table 2: List of candidate genes for the three mapped QTLs.

Acknowledgements

We would like to acknowledge project support from a Natural Sciences and Engineering Research Council of Canada (NSERC) Strategic Grant to T.B., R.B. and B.E., NSERC Discovery Grants to T.B. and B.E., and support from the FPInnovation (Montreal) fibre analysis facility at the University of British Columbia and from the Centre for Analysis of Genome Evolution and Function (CAGEF) at the University of Toronto. We also thank Claire Li, Magdalena Skocynska, Freeman Lan, Mandy Tse, Jae Kang, Hae Won Kwon, Minwoo Kim, and Andrew Kuan for their technical assistance.

References

- Alcazar R, Garcia AV, Parker JE, Reymond M.** 2009. Incremental steps toward incompatibility revealed by *Arabidopsis* epistatic interactions modulating salicylic acid pathway activation. *Proceedings of the National Academy of Sciences, USA* **106**, 334–339.
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389–3402.
- Bentsink L, Jowett J, Hanhart CJ, Koornneef M.** 2006. Cloning of *DOG1*, a quantitative trait locus controlling seed dormancy in *Arabidopsis*. *Proceedings of the National Academy of Sciences, USA* **103**, 17042–17047.
- Bentsink L, Yuan K, Koornneef M, Vreugdenhil D.** 2003. The genetics of phytate and phosphate accumulation in seeds and leaves of *Arabidopsis thaliana*, using natural variation. *Theoretical and Applied Genetics* **106**, 1234–1243.
- Bernal AJ, Yoo CM, Mutwil M, Jensen JK, Hou G, Blaukopf C, Sorensen I, Blancaflor EB, Scheller HV, Willats WG.** 2008. Functional analysis of the cellulose synthase-like genes *CSLD1*, *CSLD2*, and *CSLD4* in tip-growing *Arabidopsis* cells. *Plant Physiology* **148**, 1238–1253.
- Berthet S, Demont-Caulet N, Pollet B, et al.** 2011. Disruption of *LACCASE4* and 17 results in tissue-specific alterations to lignification of *Arabidopsis thaliana* stems. *The Plant Cell* **23**, 1124–1137.
- Biermann CJ.** 1996. *Handbook of Pulping and Papermaking*. San Diego: Academic Press.
- Bischoff V, Nita S, Neumetzler L, Schindelasch D, Urbain A, Eshed R, Persson S, Delmer D, Scheible WR.** 2010. *TRICHOME BIREFRINGENCE* and its homolog *AT5G01360* encode plant-specific DUF231 proteins required for cellulose biosynthesis in *Arabidopsis*. *Plant Physiology* **153**, 590–602.
- Bjurhager I, Olsson AM, Zhang B, Gerber L, Kumar M, Berglund LA, Burgert I, Sundberg B, Salmen L.** 2010. Ultrastructure and mechanical properties of *Populus* wood with reduced lignin content caused by transgenic down-regulation of cinnamate 4-hydroxylase. *Biomacromolecules* **11**, 2359–2365.
- Boerjan W, Ralph J, Baucher M.** 2003. Lignin biosynthesis. *Annual Review of Plant Biology* **54**, 519–546.
- Bowe LM, Coat G, dePamphilis CW.** 2000. Phylogeny of seed plants based on all three genomic compartments: extant gymnosperms are monophyletic and Gnetales' closest relatives are conifers. *Proceedings of the National Academy of Sciences, USA* **97**, 4092–4097.
- Broman KW, Wu H, Sen S, Churchill GA.** 2003. R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19**, 889–890.
- Brown DM, Goubet F, Wong VW, Goodacre R, Stephens E, Dupree P, Turner SR.** 2007. Comparison of five xylan synthesis mutants reveals new insight into the mechanisms of xylan synthesis. *The Plant Journal* **52**, 1154–1168.
- Burk DH, Ye ZH.** 2002. Alteration of oriented deposition of cellulose microfibrils by mutation of a katanin-like microtubule-severing protein. *The Plant Cell* **14**, 2145–2160.
- Byrne ME, Groover AT, Fontana JR, Martienssen RA.** 2003. Phyllotactic pattern and stem cell fate are determined by the *Arabidopsis* homeobox gene *BELLRINGER*. *Development* **130**, 3941–3950.
- Cao J, Schneeberger K, Ossowski S, et al.** 2011. Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. *Nature Genetics* **43**, 956–963.
- Cardinal AJ, Lee M, Moore KJ.** 2003. Genetic mapping and analysis of quantitative trait loci affecting fiber and lignin content in maize. *Theoretical and Applied Genetics* **106**, 866–874.
- Chang XF, Chandra R, Berleth T, Beatson RP.** 2008. Rapid, microscale, acetyl bromide-based method for high-throughput determination of lignin content in *Arabidopsis thaliana*. *Journal of Agricultural and Food Chemistry* **56**, 6825–6834.
- Chaw SM, Parkinson CL, Cheng Y, Vincent TM, Palmer JD.** 2000. Seed plant phylogeny inferred from all three plant genomes: monophyly of extant gymnosperms and origin of Gnetales from conifers. *Proceedings of the National Academy of Sciences, USA* **97**, 4086–4091.
- Chee P, Draye X, Jiang CX, Decanini L, Delmonte TA, Bredhauer R, Smith CW, Paterson AH.** 2005a. Molecular dissection of interspecific variation between *Gossypium hirsutum* and *Gossypium barbadense* (cotton) by a backcross-self approach: I. Fiber elongation. *Theoretical and Applied Genetics* **111**, 757–763.
- Chee PW, Draye X, Jiang CX, Decanini L, Delmonte TA, Bredhauer R, Smith CW, Paterson AH.** 2005b. Molecular dissection of phenotypic variation between *Gossypium hirsutum* and *Gossypium barbadense* (cotton) by a backcross-self approach: III. Fiber length. *Theoretical and Applied Genetics* **111**, 772–781.
- Chen H, Qian N, Guo W, Song Q, Li B, Deng F, Dong C, Zhang T.** 2009. Using three overlapped RILs to dissect genetically clustered QTL for fiber strength on Chro.D8 in Upland cotton. *Theoretical and Applied Genetics* **119**, 605–612.
- Coluccio MP, Sanchez SE, Kasulin L, Yanovsky MJ, Botto JF.** 2011. Genetic mapping of natural variation in a shade avoidance response: *ELF3* is the candidate gene for a QTL in hypocotyl growth regulation. *Journal of Experimental Botany* **62**, 167–176.
- Dayan J, Schwarzkopf M, Avni A, Aloni R.** 2010. Enhancing plant growth and fiber production by silencing GA 2-oxidase. *Plant Biotechnology Journal* **8**, 425–435.
- Dinwoodie JM.** 1965. The relationship between fiber morphology and paper properties: a review of literature. *TAPPI* **48**, 440–447.
- Draye X, Chee P, Jiang CX, Decanini L, Delmonte TA, Bredhauer R, Smith CW, Paterson AH.** 2005. Molecular dissection of interspecific variation between *Gossypium hirsutum* and *G. barbadense* (cotton) by a backcross-self approach: II. Fiber fineness. *Theoretical and Applied Genetics* **111**, 764–771.
- El-Din El-Assal S, Alonso-Blanco C, Peeters AJ, Raz V, Koornneef M.** 2001. A QTL for flowering time in *Arabidopsis* reveals a novel allele of *CRY2*. *Nature Genetics* **29**, 435–440.
- Eriksson ME, Israelsson M, Olsson O, Moritz T.** 2000. Increased gibberellin biosynthesis in transgenic trees promotes growth, biomass production and xylem fiber length. *Nature Biotechnology* **18**, 784–788.

- Gan X, Stegle O, Behr J, *et al.* 2011. Multiple reference genomes and transcriptomes for *Arabidopsis thaliana*. *Nature* **477**, 419–423.
- Ghandilyan A, Barboza L, Tisne S, Granier C, Reymond M, Koornneef M, Schat H, Aarts MG. 2009. Genetic analysis identifies quantitative trait loci controlling rosette mineral concentrations in *Arabidopsis thaliana* under drought. *New Phytologist* **184**, 180–192.
- Goubet F, Misrahi A, Park SK, Zhang Z, Twell D, Dupree P. 2003. AtCSLA7, a cellulose synthase-like putative glycosyltransferase, is important for pollen tube growth and embryogenesis in *Arabidopsis*. *Plant Physiology* **131**, 547–557.
- Grando S, Baum M, Ceccarelli S, Goodchild A, El-Haramein FJ, Jahoor A, Backes G. 2005. QTLs for straw quality characteristics identified in recombinant inbred lines of a *Hordeum vulgare* × *H. spontaneum* cross in a Mediterranean environment. *Theoretical and Applied Genetics* **110**, 688–695.
- Haley CS, Knott SA. 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* **69**, 315–324.
- Harada H, Kuromori T, Hirayama T, Shinozaki K, Leigh RA. 2004. Quantitative trait loci analysis of nitrate storage in *Arabidopsis* leading to an investigation of the contribution of the anion channel gene, *AtCLC-c*, to variation in nitrate levels. *Journal of Experimental Botany* **55**, 2005–2014.
- Harada H, Leigh RA. 2006. Genetic mapping of natural variation in potassium concentrations in shoots of *Arabidopsis thaliana*. *Journal of Experimental Botany* **57**, 953–960.
- Jansen RC, Stam P. 1994. High resolution of quantitative traits into multiple loci via interval mapping. *Genetics* **136**, 1447–1455.
- Jansen RC. 1993. Interval mapping of multiple quantitative trait loci. *Genetics* **135**, 205–211.
- Jansson S, Douglas CJ. 2007. *Populus*: a model system for plant biology. *Annual Review of Plant Biology* **58**, 435–458.
- Jones L, Ennos AR, Turner SR. 2001. Cloning and characterization of *irregular xylem4 (irx4)*: a severely lignin-deficient mutant of *Arabidopsis*. *The Plant Journal* **26**, 205–216.
- Kasili R, Walker JD, Simmons LA, Zhou J, De Veylder L, Larkin JC. 2010. SIAMESE cooperates with the CDH1-like protein CCS52A1 to establish endoreplication in *Arabidopsis thaliana* trichomes. *Genetics* **185**, 257–268.
- Ko JH, Yang SH, Park AH, Lerouxel O, Han KH. 2007. ANAC012, a member of the plant-specific NAC transcription factor family, negatively regulates xylary fiber development in *Arabidopsis thaliana*. *The Plant Journal* **50**, 1035–1048.
- Krakowsky MD, Lee M, Coors JG. 2005. Quantitative trait loci for cell-wall components in recombinant inbred lines of maize (*Zea mays* L.): I: stalk tissue. *Theoretical and Applied Genetics* **111**, 337–346.
- Krakowsky MD, Lee M, Coors JG. 2006. Quantitative trait loci for cell wall components in recombinant inbred lines of maize (*Zea mays* L.): II: leaf sheath tissue. *Theoretical and Applied Genetics* **112**, 717–726.
- Kramer EM. 2004. PIN and AUX/LAX proteins: their role in auxin accumulation. *Trends in Plant Science* **9**, 578–582.
- Kroymann J, Donnerhacke S, Schnabelrauch D, Mitchell-Olds T. 2003. Evolutionary dynamics of an *Arabidopsis* insect resistance quantitative trait locus. *Proceedings of the National Academy of Sciences, USA* **100** (Suppl. 2), 14587–14592.
- Lacombe E, Hawkins S, Van Doorselaere J, Piquemal J, Goffner D, Poeydomenge O, Boudet AM, Grima-Pettenati J. 1997. Cinnamoyl CoA reductase, the first committed enzyme of the lignin branch biosynthetic pathway: cloning, expression and phylogenetic relationships. *The Plant Journal* **11**, 429–441.
- Laskar DD, Jourdes M, Patten AM, Helms GL, Davin LB, Lewis NG. 2006. The *Arabidopsis* cinnamoyl CoA reductase *irx4* mutant has a delayed but coherent (normal) program of lignification. *The Plant Journal* **48**, 674–686.
- Li Q, Min D, Wang JP, Peszlen I, Horvath L, Horvath B, Nishimura Y, Jameel H, Chang HM, Chiang VL. 2011. Down-regulation of glycosyltransferase 8D genes in *Populus trichocarpa* caused reduced mechanical strength and xylan content in wood. *Tree Physiology* **31**, 226–236.
- Liepmann AH, Wightman R, Geshi N, Turner SR, Scheller HV. 2010. *Arabidopsis*—a powerful model system for plant cell wall research. *The Plant Journal* **61**, 1107–1121.
- Liepmann AH, Wilkerson CG, Keegstra K. 2005. Expression of cellulose synthase-like (Csl) genes in insect cells reveals that CslA family members encode mannan synthases. *Proceedings of the National Academy of Sciences, USA* **102**, 2221–2226.
- Lisec J, Meyer RC, Steinfath M, *et al.* 2008. Identification of metabolic and biomass QTL in *Arabidopsis thaliana* in a parallel analysis of RIL and IL populations. *The Plant Journal* **53**, 960–972.
- Lisec J, Steinfath M, Meyer RC, Selbig J, Melchinger AE, Willmitzer L, Altmann T. 2009. Identification of heterotic metabolite QTL in *Arabidopsis thaliana* RIL and IL populations. *The Plant Journal* **59**, 777–788.
- Lister C, Dean C. 1993. Recombinant inbred lines for mapping rflp and phenotypic markers in *Arabidopsis thaliana*. *The Plant Journal* **4**, 745–750.
- Mansfield SD, Kang KY, Chapple C. 2012. Designed for deconstruction—poplar trees altered in cell wall lignification improve the efficacy of bioethanol production. *New Phytol* **194**, 91–101.
- Masle J, Gilmore SR, Farquhar GD. 2005. The *ERECTA* gene regulates plant transpiration efficiency in *Arabidopsis*. *Nature* **436**, 866–870.
- Mattsson J, Sung ZR, Berleth T. 1999. Responses of plant vascular systems to auxin transport inhibition. *Development* **126**, 2979–2991.
- Mayer AM, Staples RC. 2002. Laccase: new functions for an old enzyme. *Phytochemistry* **60**, 551–565.
- McKay JK, Richards JH, Nemali KS, Sen S, Mitchell-Olds T, Boles S, Stahl EA, Wayne T, Juenger TE. 2008. Genetics of drought adaptation in *Arabidopsis thaliana* II. QTL analysis of a new mapping population, KAS-1 × TSU-1. *Evolution* **62**, 3014–3026.
- Mehrshahi P, Gonzalez-Jorge S, Akhtar TA, *et al.* 2010. Functional analysis of folate polyglutamylation and its essential role in plant metabolism and development. *The Plant Journal* **64**, 267–279.
- Mei M, Syed NH, Gao W, Thaxton PM, Smith CW, Stelly DM, Chen ZJ. 2004. Genetic mapping and QTL analysis of fiber-related traits in cotton (*Gossypium*). *Theoretical and Applied Genetics* **108**, 280–291.

- Messerschmidt A, Huber R.** 1990. The blue oxidases, ascorbate oxidase, laccase and ceruloplasmin. Modelling and structural relationships. *European Journal of Biochemistry* **187**, 341–352.
- Mizrachi E, Mansfield SD, Myburg AA.** 2012. Cellulose factories: advancing bioenergy production from forest trees. *New Phytologist* **194**, 54–62.
- Morrison IM.** 1972. A semi-micro method for the determination of lignin and its use in predicting the digestibility of forage crops. *Journal of the Science of Food and Agriculture* **23**, 791.
- Mouchel CF, Briggs GC, Hardtke CS.** 2004. Natural genetic variation in *Arabidopsis* identifies *BREVIS RADIX*, a novel regulator of cell proliferation and elongation in the root. *Genes & Development* **18**, 700–714.
- Nieminen K, Robischon M, Immanen J, Helariutta Y.** 2012. Towards optimizing wood development in bioenergy trees. *New Phytologist* **194**, 46–53.
- Odendahl S.** 1994. Environmental protection and consumer demands: a review of trends and impacts. *Pulp and Paper Canada* **95**, 144–148.
- Olson JA, Robertson AG, Finnigan TD, Turner RRH.** 1995. An analyzer for fiber length and shape. *Journal of Pulp and Paper Science* **21**, J376.
- Page DH, Seth RS.** 1980. The elastic modulus of paper III. the effect of dislocations, microcompressions, curl and kinks. *TAPPI* **63**, 99–102.
- Park S, Rancour DM, Bednarek SY.** 2008. In planta analysis of the cell cycle-dependent localization of AtCDC48A and its critical roles in cell division, expansion, and differentiation. *Plant Physiology* **148**, 246–258.
- Paterson AH.** 2011. QTL alleles for improved fiber quality from a wild Hawaiian cotton, *Gossypium tomentosum*. *Theoretical and Applied Genetics* **123**, 1075–1088.
- Patten AM, Cardenas CL, Cochrane FC, Laskar DD, Bedgar DL, Davin LB, Lewis NG.** 2005. Reassessment of effects on lignification and vascular development in the *irx4* *Arabidopsis* mutant. *Phytochemistry* **66**, 2092–2107.
- Peaucelle A, Louvet R, Johansen JN, et al.** 2011. The transcription factor BELLRINGER modulates phyllotaxis by regulating the expression of a pectin methylesterase in *Arabidopsis*. *Development* **138**, 4733–4741.
- Pfalz M, Vogel H, Mitchell-Olds T, Kroymann J.** 2007. Mapping of QTL for resistance against the crucifer specialist herbivore *Pieris brassicae* in a new *Arabidopsis* inbred line population, Da(1)-12×EI-2. *PLoS One* **2**, e578.
- Pilate G, Guiney E, Holt K, et al.** 2002. Field and pulping performances of transgenic trees with altered lignification. *Nature Biotechnology* **20**, 607–612.
- Qin H, Guo W, Zhang YM, Zhang T.** 2008. QTL mapping of yield and fiber traits based on a four-way cross population in *Gossypium hirsutum* L. *Theoretical and Applied Genetics* **117**, 883–894.
- R Development Core Team. 2009. *R: A Language and Environment for Statistical Computing*. Vienna, Austria.
- Ragni L, Nieminen K, Pacheco-Villalobos D, Sibout R, Schwechheimer C, Hardtke CS.** 2011. Mobile gibberellin directly stimulates *Arabidopsis* hypocotyl xylem expansion. *The Plant Cell* **23**, 1322–1336.
- Ralph J, Hatfield RD, Piquemal J, Yahiaoui N, Pean M, Lapierre C, Boudet AM.** 1998. NMR characterization of altered lignins extracted from tobacco plants down-regulated for lignification enzymes cinnamylalcohol dehydrogenase and cinnamoyl-CoA reductase. *Proceedings of the National Academy of Sciences, USA* **95**, 12803–12808.
- Rancour DM, Dickey CE, Park S, Bednarek SY.** 2002. Characterization of AtCDC48. Evidence for multiple membrane fusion mechanisms at the plane of cell division in plants. *Plant Physiology* **130**, 1241–1253.
- Ranocha P, Chabannes M, Chamayou S, Danoun S, Jauneau A, Boudet AM, Goffner D.** 2002. Laccase down-regulation causes alterations in phenolic metabolism and cell wall structure in poplar. *Plant Physiology* **129**, 145–155.
- Ranocha P, Denance N, Vanholme R, et al.** 2010. *Walls are thin 1* (*WAT1*), an *Arabidopsis* homolog of *Medicago truncatula* *NODULIN21*, is a tonoplast-localized protein required for secondary wall formation in fibers. *The Plant Journal* **63**, 469–483.
- Reiter RS, Williams JG, Feldmann KA, Rafalski JA, Tingey SV, Scolnik PA.** 1992. Global and local genome mapping in *Arabidopsis thaliana* by using recombinant inbred lines and random amplified polymorphic DNAs. *Proceedings of the National Academy of Sciences, USA* **89**, 1477–1481.
- Reymond M, Svistoonoff S, Loudet O, Nussaume L, Desnos T.** 2006. Identification of QTL controlling root growth response to phosphate starvation in *Arabidopsis thaliana*. *Plant, Cell & Environment* **29**, 115–125.
- Roberston AG, Olson JA, Allen PA, Chan B, Seth RS.** 1999. Measurement of fiber length, coarseness and shape with the fiber quality analyzer. *TAPPI* **82**, 30–41.
- Sanyal A, Randal Linder C.** 2011. Quantitative trait loci involved in regulating seed oil composition in *Arabidopsis thaliana* and their evolutionary implications. *Theoretical and Applied Genetics* **124**, 723–738.
- Schillmiller AL, Stout J, Weng JK, Humphreys J, Ruegger MO, Chapple C.** 2009. Mutations in the *cinnamate 4-hydroxylase* gene impact metabolism, growth and development in *Arabidopsis*. *The Plant Journal* **60**, 771–782.
- Sen S, Churchill GA.** 2001. A statistical framework for quantitative trait mapping. *Genetics* **159**, 371–387.
- Sewell MM, Davis MF, Tuskan GA, Wheeler NC, Elam CC, Bassoni DL, Neale DB.** 2002. Identification of QTLs influencing wood property traits in loblolly pine (*Pinus taeda* L.). II. Chemical wood properties. *Theoretical and Applied Genetics* **104**, 214–222.
- Shiringani AL, Friedt W.** 2011. QTL for fibre-related traits in grain × sweet sorghum as a tool for the enhancement of sorghum as a biomass crop. *Theoretical and Applied Genetics* **123**, 999–1011.
- Singer T, Fan Y, Chang HS, Zhu T, Hazen SP, Briggs SP.** 2006. A high-resolution map of *Arabidopsis* recombinant inbred lines by whole-genome exon array hybridization. *PLoS Genetics* **2**, e144.
- Smith HM, Hake S.** 2003. The interaction of two homeobox genes, *BREVIPEDICELLUS* and *PENNYWISE*, regulates internode

- patterning in the *Arabidopsis* inflorescence. *The Plant Cell* **15**, 1717–1727.
- Soltis DE, Soltis PS, Zanis MJ.** 2002. Phylogeny of seed plants based on evidence from eight genes. *American Journal of Botany* **89**, 1670–1681.
- Somerville C, Youngs H, Taylor C, Davis SC, Long SP.** 2010. Feedstocks for lignocellulosic biofuels. *Science* **329**, 790–792.
- Swarup K, Alonso-Blanco C, Lynn JR, Michaels SD, Amasino RM, Koornneef M, Millar AJ.** 1999. Natural allelic variation identifies new genes in the *Arabidopsis* circadian system. *The Plant Journal* **20**, 67–77.
- TAPPI. 1992. *Fines fraction of paper stock by wet screening. Test Method T 261 cm-90*. Atlanta, Georgia, USA: TAPPI Press.
- Torii KU, Mitsukawa N, Oosumi T, Matsuura Y, Yokoyama R, Whittier RF, Komeda Y.** 1996. The *Arabidopsis* *ERECTA* gene encodes a putative receptor protein kinase with extracellular leucine-rich repeats. *The Plant Cell* **8**, 735–746.
- Toufighi K, Brady SM, Austin R, Ly E, Provart NJ.** 2005. The Botany Array Resource: e-Northern, Expression Angling, and promoter analyses. *The Plant Journal* **43**, 153–163.
- Tuskan GA, Difazio S, Jansson S, et al.** 2006. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **313**, 1596–1604.
- Ulloa M, Saha S, Jenkins JN, Meredith WR Jr, McCarty JC Jr, Stelly DM.** 2005. Chromosomal assignment of RFLP linkage groups harboring important QTLs on an intraspecific cotton (*Gossypium hirsutum* L.) Joinmap. *Journal of Heredity* **96**, 132–144.
- Van Leene J, Hollunder J, Eeckhout D, et al.** 2010. Targeted interactomics reveals a complex core cell cycle machinery in *Arabidopsis thaliana*. *Molecular Systems Biology* **6**, 397.
- Via BK, Stine M, Shupe TF, So CL, Groom L.** 2004. Genetic improvement of fiber length and coarseness based on paper product performance and material variability: a review. *IAWA Journal* **25**, 401–414.
- Walker JD, Oppenheimer DG, Conciennne J, Larkin JC.** 2000. *SIAMESE*, a gene controlling the endoreduplication cell cycle in *Arabidopsis thaliana* trichomes. *Development* **127**, 3931–3940.
- Waters BM, Grusak MA.** 2008. Quantitative trait locus mapping for seed mineral concentrations in two *Arabidopsis thaliana* recombinant inbred populations. *New Phytologist* **179**, 1033–1047.
- Werner JD, Borevitz JO, Warthmann N, Trainer GT, Ecker JR, Chory J, Weigel D.** 2005. Quantitative trait locus mapping and DNA array hybridization identify an FLM deletion as a cause for natural flowering-time variation. *Proceedings of the National Academy of Sciences, USA* **102**, 2460–2465.
- Wienkoop S, Baginsky S, Weckwerth W.** 2010. *Arabidopsis thaliana* as a model organism for plant proteome research. *Journal of Proteomics* **73**, 2239–2248.
- Wu G, Otegui MS, Spalding EP.** 2010. The ER-localized TWD1 immunophilin is necessary for localization of multidrug resistance-like proteins required for polar auxin transport in *Arabidopsis* roots. *The Plant Cell* **22**, 3295–3304.
- Xie JK, Kong XL, Chen J, Hu BL, Wen P, Zhuang JY, Bao JS.** 2011. Mapping of quantitative trait loci for fiber and lignin contents from an interspecific cross *Oryza sativa* × *Oryza rufipogon*. *Journal of Zhejiang University Science B* **12**, 518–526.
- Yamaguchi M, Mitsuda N, Ohtani M, Ohme-Takagi M, Kato K, Demura T.** 2011. VASCULAR-RELATED NAC-DOMAIN7 directly regulates the expression of a broad range of genes for xylem vessel formation. *The Plant Journal* **66**, 579–590.
- Yonekura-Sakakibara K, Tohge T, Matsuda F, Nakabayashi R, Takayama H, Niida R, Watanabe-Takahashi A, Inoue E, Saito K.** 2008. Comprehensive flavonol profiling and transcriptome coexpression analysis leading to decoding gene-metabolite correlations in *Arabidopsis*. *The Plant Cell* **20**, 2160–2176.
- Yu Q, Pulkkinen P, Rautio M, Haapanen M, Alén R, Stener LG, Beuker E, Tigerstedt PM.** 2001. Genetic control of wood physicochemical properties, growth, and phenology in hybrid aspen clones. *Canadian Journal of Forest Research* **31**, 1348–1356.
- Zawaski C, Kadmiel M, Ma C, Gai Y, Jiang X, Strauss SH, Busov VB.** 2011. *SHORT INTERNODES-like* genes regulate shoot growth and xylem proliferation in *Populus*. *New Phytologist* **191**, 678–691.
- Zeng C, Han Y, Shi L, Peng L, Wang Y, Xu F, Meng J.** 2008. Genetic analysis of the physiological responses to low boron stress in *Arabidopsis thaliana*. *Plant, Cell & Environment* **31**, 112–122.
- Zeng ZB.** 1993. Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. *Proceedings of the National Academy of Sciences, USA* **90**, 10972–10976.
- Zeng ZB.** 1994. Precision mapping of quantitative trait loci. *Genetics* **136**, 1457–1468.
- Zhang C, Mallery EL, Schlueter J, Huang S, Fan Y, Brankle S, Staiger CJ, Szymanski DB.** 2008. *Arabidopsis* SCARs function interchangeably to meet actin-related protein 2/3 activation thresholds during morphogenesis. *The Plant Cell* **20**, 995–1011.
- Zhang Z, Ober JA, Kliebenstein DJ.** 2006. The gene controlling the quantitative trait locus *EPITHIOSPECIFIER MODIFIER1* alters glucosinolate hydrolysis and insect resistance in *Arabidopsis*. *The Plant Cell* **18**, 1524–1536.
- Zhang Z, Rong J, Waghmare VN, Chee PW, May OL, Wright RJ, Gannaway JR, Chew YH, Halliday KJ.** 2011. A stress-free walk from *Arabidopsis* to crops. *Current Opinion in Biotechnology* **22**, 281–286.
- Zhong R, Pena MJ, Zhou GK, et al.** 2005. *Arabidopsis fragile fiber8*, which encodes a putative glucuronyltransferase, is essential for normal secondary wall synthesis. *The Plant Cell* **17**, 3390–3408.
- Zhong R, Taylor JJ, Ye ZH.** 1997. Disruption of interfascicular fiber differentiation in an *Arabidopsis* mutant. *The Plant Cell* **9**, 2159–2170.
- Zhong R, Ye ZH.** 2001. Alteration of auxin polar transport in the *Arabidopsis* *ifl1* mutants. *Plant Physiology* **126**, 549–563.